Pareto efficiency for the concave order and multivariate comonotonicity

Guillaume Carlier, Rose-Anne Dana, Alfred Galichon

Pareto efficiency for the concave order and multivariate comonotonicity

G. Carlier [*], R.-A. Dana [†], A. Galichon [‡]

October 12, 2011

Abstract

This paper studies efficient risk-sharing rules for the concave dominance order. For a univariate risk, it follows from a *comonotone dominance principle*, due to Landsberger and Meilijson [27], that efficiency is characterized by a comonotonicity condition. The goal of the paper is to generalize the comonotone dominance principle as well as the equivalence between efficiency and comonotonicity to the multidimensional case. The multivariate case is more involved (in particular because there is no immediate extension of the notion of comonotonicity), and it is addressed by using techniques from convex duality and optimal transportation.

JEL classification: C61, D61, D81.

Keywords: concave order, stochastic dominance, comonotonicity, efficiency, multivariate risk-sharing.

1 Introduction

Motivation. The aim of this paper is to study Pareto efficient allocations of risky consumptions of multiple goods in a contingent exchange economy.

[*]CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE carlier@ceremade.dauphine.fr

[†]CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE dana@ceremade.dauphine.fr

[‡]Département d'Economie, Ecole Polytechnique, alfred.galichon@polytechnique.edu. Galichon gratefully acknowledges support from Chaire EDF-Calyon "Finance and Développement Durable," Chaire FDIR "Socially Responsible Investments and Value Creation" and Chaire Axa "Assurance et Risques Majeurs" and FiME, Laboratoire de Finance des Marchés de l'Energie (www.fime-lab.org).

In this framework, consumption goods are imperfect substitutes, hence consumption is measured along several different units, instead of being denominated in one single monetary value. These units can be for instance material consumption and labor, or future consumptions at various subsequent dates, or currency units with limited exchangeability. In this setting, risky consumption can no longer be represented as a *random variable*, but as a *random vector*.

Agents are assumed to have incomplete preferences associated with the *concave order*: a risk (random vector) X is preferred to a risk Y in the concave order whenever *every* risk-averse expected utility decision-maker prefers X to Y. Again, these preferences form an incomplete order, hence this assumption (and its empirical content) may appear as relatively weak. We shall see, however, that they in fact lead to strong predictions.

The motivation of the paper is to characterize efficient allocations for the concave order on observable data (for instance, insurance contracts). In the case of univariate risk, it is known that efficiency for the concave order is equivalent to efficiency for some strictly concave expected utility model, which in turn yields a tractable characterization of efficiency: *comonotonicity* of the allocations. Allocations are comonotone whenever each agent's contingent consumption is a nondecreasing function of the aggregate consumption. Further, a *comonotone dominance principle* can also be proven: if some initial allocation is not comonotone, there is a comonotone allocation such that every agent weakly prefers their contingent consumption in the new allocation (at least one preferring strictly). Comonotonicity fully characterizes efficiency and it is a testable and tractable property. Moreover, as a consequence of the comonotone dominance principle, attention may be restricted to the set of comonotone allocations, which is convex and almost compact. Hence existence results may be obtained for many risk-sharing problems (see for instance [26] in the framework of risk measures, or [9], [10] for classes of law invariant and concave utilities).

Main results. This paper is devoted to the extension of the comonotone dominance result and its application to the characterization of efficient allocations in the multivariate setting. To this end, a definition of comonotonicity for the multivariate case is first needed. Roughly speaking, according to the definition of multivariate comonotonicity we adopt, an allocation $(X_1, ..., X_p)$ (with each X_i being random vectors) is comonotone if it is efficient for some strictly concave expected utility model. By first order conditions, this implies that there is a random vector Z and convex functions φ_i such that $X_i = \nabla\varphi_i(Z)$, where $\nabla\varphi_i$ is the gradient map of φ_i. This is the definition of multivariate comonotonicity used by Ekeland, Galichon and Henry in [20].

The comonotone dominance principle is next extended by solving a variational problem. More precisely, given an initial allocation and a collection of strictly concave utility functions, we maximize the sum of these utilities among allocations that dominate the initial allocation. We prove, and this is the hard part of the proof, that the corresponding optimal allocation is necessarily comonotone. The precise statement of the multivariate comonotone dominance result is, however more complicated than in the univariate case since it requires the use of weak closures and a concept slightly stronger than strict convexity. This follows from the fact that the set of multivariate comonotone allocations is neither convex nor compact (even up to constants), contrary to the univariate case (counterexamples are given). Finally one may wonder whether the equivalence between efficient and comonotone allocations is preserved in the multivariate case. The answer is yes, up to some (interesting) technicalities. Again the precise statement of the result is more complicated than in the univariate case. When applied to the univariate case, our proof of the comonotone dominance result improves upon all existing proofs (see [27], [14] and [30]). Indeed, it addresses directly the case of many agents, it uses neither the discrete case nor a limiting argument, and no hypotheses need be made on the aggregate endowment.

Literature overview. There is a distinguished tradition in modeling preferences by concave dominance. Introduced in economics by Rothschild and Stiglitz [34], the concave order has then been used in a wide variety of economic contexts. To give a few references, let us mention efficiency pricing (Peleg and Yaari [32], Chew and Zilcha [12]), measurement of inequality (Atkinson [3]), and finance (Dybvig [16], Jouini and Kallal [25]).

In dimension one, the mutuality principle arose in the early work of Borch [6], Arrow [1], [2] and Wilson [38]; see also LeRoy and Werner [28]. Landsberger and Meilijson [27] proved (for two agents and a discrete setting) that any allocation of a given aggregate risk is dominated in the sense of concave dominance by a comonotone allocation. This comonotone dominance principle has been extended to the continuous case by limiting arguments (see [14] and [30]). It implies the comonotonicity of efficient allocations for the concave order. The equivalence between comonotonicity and efficiency was only proved recently by Dana [13] for the discrete case and by Dana and Meilijson [14] for the continuous case. This equivalence stimulated a line of research on comonotonicity in the insurance and finance literature, see for instance Jouini and Napp [23], [24]. On the empirical side, Townsend [37] proposed to test whether the mutuality principle holds in three poor villages in southern India while Attanasio and Davis ([4]) worked with US labor data. The general findings of these empirical studies is that comonotonicity can be usually strongly rejected. A possible explanation of why efficiency is usually not observed in the data is that the aforementioned literature only considers risk-sharing in the case of one good (monetary consumption) and does not take into account the cross-subsidy effects between several risky goods which are only imperfect substitutes. Other papers, such as Brown and Matzkin ([8]) have tried to test whether observed market data on prices, aggregate endowments and individual incomes satisfy the restrictions that are imposed by Walrasian equilibrium. In contrast to this approach, we do not assume prices to be available to the researcher.

The notion of multivariate comonotonicity adopted in this paper coincides (up to some technical details) with the one originally introduced by Ekeland, Galichon and Henry in [20] under the name μ-comonotonicity, in the context of risk measures. Galichon and Henry use that concept to generalize rank-dependent expected utility in [19]. Other proposals for multivariate comonotonicity exist and are reviewed e.g. in [33]; however they do not seem to be related to efficient risk-sharing. While the results of [20] are strongly related to maximal correlation functionals and to the quadratic optimal transportation problem (and in particular Brenier's seminal paper [7]), the present approach will rely on a slightly different optimization problem that has some familiarities with the multi-marginals optimal transport problem of Gangbo and Święch [22].

Organization of the paper. The paper is organized as follows. Section 2 recalls some definitions and various characterizations of comonotonicity in the univariate case. Section 3 revisits the comonotone dominance principle of [27] and characterizes efficient risk sharing in the univariate case. A notion of multivariate comonotonicity is introduced in Section 4, an analogue of the comonotone dominance principle is stated, and efficient sharing-rules are characterized as the weak closure of comonotone allocations. Section 5 concludes the paper. Proofs are gathered in section 6.

2 Preliminaries

Given as primitive is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For every (univariate or multivariate) random vector X on such space, the law of X is denoted $\mathcal{L}(X)$. Two random vectors X and Y are called equivalent in distribution (denoted $X \sim Y$), if $\mathcal{L}(X) = \mathcal{L}(Y)$.

Definition 2.1. *Let X and Y be bounded random vectors with values in $\mathbb{R}^d$, then X dominates Y for the concave order, denoted $X \succcurlyeq Y$, if and only if $\mathbb{E}(\varphi(X)) \leq \mathbb{E}(\varphi(Y))$ for every convex function $\varphi : \mathbb{R}^d \to \mathbb{R}$. If, in addition,*

$\mathbb{E}(\varphi(X)) < \mathbb{E}(\varphi(Y))$ *for some convex function* φ*, then* X *is said to dominate* Y *strictly.*

As the paper makes extensive use of *convex* analysis (Legendre transforms, infimal convolutions, convex duality), the concave order is defined here in terms of convex loss functions while usually defined with concave utilities. Clearly the definition above coincides with the standard one. As $X \succcurlyeq Y$ implies that $\mathbb{E}(X) = \mathbb{E}(Y)$, comparing risks for $\succcurlyeq$ only makes sense for random vectors with the same mean. We refer to Rothschild and Stiglitz [34] and Föllmer and Schied [21] for various characterizations of concave dominance in the univariate case and to Müller and Stoyan [31] for the multivariate case. Using a classical result of Cartier, Fell and Meyer (see [11] or [36]), one deduces a convenient characterization (see section 6 for a proof) of strict dominance as follows:

Lemma 2.2. *Let* X *and* Y *be bounded random vectors with values in* $\mathbb{R}^d$*, then the following statements are equivalent:*

1. X *strictly dominates* Y*,*
2. $X \succcurlyeq Y$ *and* $\mathcal{L}(X) \neq \mathcal{L}(Y)$*,*
3. $X \succcurlyeq Y$ *and for every strictly convex function* φ*,* $\mathbb{E}(\varphi(X)) < \mathbb{E}(\varphi(Y))$*.*

Given $X \in L^\infty(\Omega, \mathbb{R}^d)$, a random vector of aggregate risk of dimension $d \geq 1$, the set of admissible allocations or risk-sharing of X among p agents is denoted $\mathcal{A}(X)$:

$$\mathcal{A}(X) := \{\mathbf{Y} = (Y_1, ..., Y_p) \in L^\infty(\Omega, \mathbb{R}^d) : \sum_{i=1}^p Y_i = X\}.$$

For simplicity, the dependence of $\mathcal{A}(X)$ on the number p of agents does not appear explicitly. A concept of dominance for allocations of X is defined next.

Definition 2.3. *For* $d \geq 1$*, let* $\mathbf{X} = (X_1, ..., X_p)$ *and* $\mathbf{Y} := (Y_1, ..., Y_p)$ *be in* $\mathcal{A}(X)$*. Then* $\mathbf{X}$ *is said to dominate* $\mathbf{Y}$ *if* $X_i \succcurlyeq Y_i$ *for every* $i \in \{1, ..., p\}$*. If, in addition, there is an* $i \in \{1, ..., p\}$ *such that* X_i *strictly dominates* Y_i*, then* $\mathbf{X}$ *is said to strictly dominate* $\mathbf{Y}$*. An allocation* $\mathbf{X} \in \mathcal{A}(X)$ *is Pareto-efficient (for the concave order) if there is no allocation in* $\mathcal{A}(X)$ *that strictly dominates* $\mathbf{X}$*.*

It may easily be verified that dominance of allocations can also be defined as follows. Let $\mathbf{X}$ and $\mathbf{Y}$ be in $\mathcal{A}(X)$, then $\mathbf{X}$ dominates $\mathbf{Y}$ if and only if

$$\mathbb{E}(\sum_{i=1}^{p} \varphi_i(X_i)) \leq \mathbb{E}(\sum_{i=1}^{p} \varphi_i(Y_i)), \tag{2.1}$$

for every collection of convex functions $\varphi_i : \mathbb{R}^d \to \mathbb{R}$. Moreover, $\mathbf{X}$ strictly dominates $\mathbf{Y}$ if and only if the previous inequality is strict for some collection of convex functions $\varphi_i : \mathbb{R}^d \to \mathbb{R}$. Note that from lemma 2.2, it is equivalent to require that the inequality is strict for every collection of strictly convex functions. Therefore, if $\mathbf{X}$ is the solution of the problem

$$\inf\Big\{\sum_{i=1}^{p} \mathbb{E}(\varphi_i(Y_i)) \ : \ (Y_1, ..., Y_p) \in \mathcal{A}(X)\Big\} \tag{2.2}$$

for some collection of strictly convex functions φ_i, then $\mathbf{X}$ is efficient. Finally, recall that in the univariate case, comonotonicity is defined by:

Definition 2.4. *A collection $(X_1, ..., X_p)$ of p real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ is comonotone if for every $(i, j) \in \{1, ..., p\}^2$,*

$$(X_i(\omega') - X_i(\omega))(X_j(\omega') - X_j(\omega)) \geq 0 \text{ for } \mathbb{P} \otimes \mathbb{P}\text{-a.e. } (\omega, \omega') \in \Omega^2.$$

It is well-known that comonotonicity of $(X_1, ..., X_p)$ is equivalent to the fact that each X_i can be written as a nondecreasing function of the sum $X = \sum_i X_i$ (see for instance Denneberg [15]). Therefore $(X_1, ..., X_p)$ is comonotone if and only if there are nondecreasing functions f_i summing to the identity such that $X_i = f_i(X)$. Note that the functions f_i are all 1-Lipschitz . The extension of this notion to the multivariate case (i.e when each X_i is $\mathbb{R}^d$-valued) is not immediately obvious and will be addressed in Section 4.

We now provide another characterization of comonotonicity based on the notion of maximal correlation. From now on, assume that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic which means that there is no $A \in \mathcal{F}$ such that for every $B \in \mathcal{F}$ if $\mathbb{P}(B) < \mathbb{P}(A)$ then $\mathbb{P}(B) = 0$. It is well-known that $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic if and only if a random variable that is uniformly distributed on $[0, 1]$, which is denoted $U \sim \mathcal{U}([0, 1])$, can be constructed on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, and define for every $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, both Z and X being univariate here, the maximal correlation functional:

$$\varrho_Z(X) := \sup_{\tilde{X} \sim X} \mathbb{E}(Z\tilde{X}) = \sup_{\tilde{Z} \sim Z} \mathbb{E}(\tilde{Z}X) = \sup_{\tilde{Z} \sim Z, \, \tilde{X} \sim X} \mathbb{E}(\tilde{Z}\tilde{X}). \tag{2.3}$$

The functional ϱ_Z has extensively been discussed in economics and in finance, therefore only a few useful facts are recalled. Let F_X^{-1} be the quantile function of X, that is the pseudo-inverse of distribution function F_X. From Hardy-Littlewood's inequality, one has

$$\varrho_Z(X) = \int_0^1 F_X^{-1}(t) F_Z^{-1}(t) dt,$$

and the supremum in (2.3) is achieved by any pair $(\tilde{Z}, \tilde{X})$ of comonotone random variables $(F_Z^{-1}(U), F_X^{-1}(U))$ for U uniformly distributed. By symmetry, one can either fix Z or fix X. Fixing for instance Z, the supremum is achieved by $F_X^{-1}(U)$ where $U \sim \mathcal{U}([0,1])$ and $Z = F_Z^{-1}(U)$. When Z is non-atomic, there is a unique $U = F_Z(Z)$ such that $Z = F_Z^{-1}(U)$, and the supremum is uniquely attained by the non-decreasing function of Z, $F_X^{-1} \circ F_Z(Z)$:

$$\varrho_Z(X) = \mathbb{E}(Z F_X^{-1} \circ F_Z(Z)). \tag{2.4}$$

Also note that ϱ_Z is subadditive: $\varrho_Z(\sum_i X_i) \leq \sum_i \varrho_Z(X_i)$.

Proposition 2.5. *Let $(X_1, ..., X_p)$ be in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. The following assertions are equivalent:*

1. *$(X_1, ..., X_p)$ are comonotone,*
2. *for any non-atomic $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$,*

$$\varrho_Z\left(\sum_i X_i\right) = \sum_i \varrho_Z(X_i), \tag{2.5}$$

3. *for some non-atomic $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, (2.5) holds true.*

Proof. For the sake of simplicity, we restrict ourselves to $p = 2$ and set $(X_1, X_2) = (X, Y)$. Point 1 implies point 2 since $F_{X+Y}^{-1} = F_X^{-1} + F_Y^{-1}$ for comonotone X and Y. To show that point 3 implies point 1, assume that for some non-atomic Z, one has (2.5), which by sublinearity is equivalent to $\varrho_Z(X+Y) \geq \varrho_Z(X) + \varrho_Z(Y)$. Let Z_{X+Y} (resp. Z_X and Z_Y) be distributed as Z and solve $\sup_{\tilde{Z} \sim Z} E(\tilde{Z} X)$ (resp. $\varrho_Z(X)$ and $\varrho_Z(Y)$). One then has:

$$\mathbb{E}(Z_{X+Y}(X+Y)) \geq \mathbb{E}(Z_X X) + \mathbb{E}(Z_Y Y).$$

As $\mathbb{E}(Z_{X+Y} X) \leq \mathbb{E}(Z_X X)$ and $\mathbb{E}(Z_{X+Y} Y) \leq \mathbb{E}(Z_Y Y)$, it follows $\mathbb{E}(Z_{X+Y} X) = \mathbb{E}(Z_X X) = \varrho_Z(X)$ and $\mathbb{E}(Z_{X+Y} Y) = \mathbb{E}(Z_Y Y) = \varrho_Z(Y)$, hence from (2.4), $X = F_X^{-1} \circ F_{Z_{X+Y}}(Z_{X+Y}))$ and $Y = F_Y^{-1} \circ F_{Z_{X+Y}}(Z_{X+Y}))$, proving comonotonicity. $\square$

Proposition 2.5 was the starting point of Ekeland, Galichon and Henry [20] for providing a multivariate generalization of the concept of comonotonicity. In the sequel we shall further discuss this multivariate extension and compare it with the one proposed in the present paper.

3 The univariate case

A landmark result, due to Landsberger and Meilijson [27] states that any allocation is dominated by a comonotone one. The original proof was given in the discrete case for two agents, and the results were extended to the general case by approximation. We give an alternative proof in the Appendix based on the same approach we shall use in the multidimensional case. This proof is based on a certain optimization problem; we believe that, even in the unidimensional case, it is of interest per se since it does not require approximation arguments and slightly improves on the original statement by proving strict dominance of non-comonotone allocations. Contrary to Landsberger and Meilijson, one needs however to assume, as before, that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic.

Theorem 3.1. *Let X be a bounded real-valued random variable on the non-atomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ be an allocation. There exists a comonotone allocation in $\mathcal{A}(X)$ that dominates $\mathbf{X}$. Moreover, if $\mathbf{X}$ is not comonotone, then there exists an allocation that strictly dominates $\mathbf{X}$.*

As an application, we have:

Theorem 3.2. *Let X be a bounded real-valued random variable on the non-atomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$. Then the following statements are equivalent:*

1. *$\mathbf{X}$ is efficient,*
2. *$\mathbf{X}$ is comonotone,*
3. *there exist continuous and strictly convex functions $(\psi_1, ..., \psi_p)$ such that $\mathbf{X}$ solves*

$$\inf\{\sum_{i=1}^{p} \mathbb{E}(\psi_i(Y_i)) \ : \ \sum_{i=1}^{p} Y_i = X\}.$$

Proof. Point 1 implies point 2: the comonotonicity of efficient allocations of X follows directly from Theorem 3.1. Point 2 implies point 3: if $\mathbf{X} = (X_1, ..., X_p)$ is comonotone, let us write $X_i = f_i(X)$ for some nondecreasing and 1-Lipschitz functions f_i: $[m, M] \to \mathbb{R}$ (with $M := \mathrm{Esssup} X$, $m := \mathrm{Essinf} X$) summing up to the identity map. Extending the f_i functions by $f_i(x) = f_i(M) + (x - M)/p$ for $x \geq M$ and $f_i(x) = f_i(m) + (x - m)/p$ for $x \leq m$, one gets 1-Lipschitz nondecreasing functions summing up to the identity everywhere. Let $\varphi(x) := \int_0^x f_i(s)ds$ for every x. The functions φ_i are convex and $C^{1,1}$ (i.e. C^1 with a Lipschitz continuous derivative) and have quadratic growth at ∞. The convex conjugates[1] $\psi_i := \varphi_i^*$ are strictly convex and continuous functions, and by construction, one has for every i, $X \in \partial\psi_i(X_i)$ a.s., which implies that $(X_1, ..., X_p)$ minimizes $\mathbb{E}(\sum_i \psi_i(Y_i))$ subject to $\sum_i Y_i = X$, which proves point 3. Point 3 implies point 1 since the ψ_i functions are strictly convex; if $(X_1, ..., X_p)$ satisfies point 3 then it is an efficient allocation of X.

□

Corollary 3.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be non-atomic, then the set of efficient allocations of X is convex and compact in L^∞ up to zero-sum translations (which means that it can be written as $\{(\lambda_1, ..., \lambda_p) \ : \ \sum_{i=1}^p \lambda_i = 0\} + A_0$ with A_0 compact in L^∞). In particular, the set of efficient allocations of X is closed in L^∞.*

Proof. Let $M := \mathrm{Esssup} X$, $m := \mathrm{Essinf} X$ and define K_0 as the set of functions $(f_1, ..., f_p) \in C([m, M], \mathbb{R}^p)$ such that for each nondecreasing f_i, $f_i(m) = m/p$ and $\sum_{i=1}^p f_i(x) = x$ for every $x \in [m, M]$, and let

$$K := K_0 + \{(\lambda_1, ..., \lambda_p) \ : \ \sum_{i=1}^p \lambda_i = 0\}.$$

The convexity claim thus follows from theorem 3.2 and the convexity of K. Let us remark that elements of K_0 have 1-Lipschitz components and are bounded. The compactness of K in $C([m, M], \mathbb{R}^p)$ then follows from Ascoli's theorem. The compactness and closedness claims directly follow. □

Convexity and compactness of efficient allocations are quite remarkable features and as will be shown later, they are no longer true in the multivariate case. Note also that efficient allocations are regular: they are 1-Lipschitz functions of aggregate risk.

[1] Let us recall that the Legendre transform or convex conjugate of φ_i is by definition given by $\varphi_i^*(x) := \sup_y \{x \cdot y - \varphi_i(y)\}$.

4 The multivariate case

The aim of this section is to generalize to the multivariate case the results obtained in the univariate case. More particularly, Landsberger and Meilijson's comonotone dominance principle are extended: 1) any allocation is dominated by a comonotone allocation; 2) any non comonotone allocation is strictly dominated by a comonotone one.

When addressing these generalizations it is not immediately clear what is the appropriate notion of comonotonicity in the multivariate framework. Let us informally give an intuitive presentation of the approach developed in the following paragraphs. A natural generalization of monotone maps in several dimensions is given by subgradients of convex functions. It is therefore tempting to say that an allocation $(X_1, ..., X_p) \in \mathcal{A}(X)$ is comonotone whenever there is a common random vector Z (interpreted as a price) and convex functions φ_i (interpreted as individual costs) such that $X_i \in \partial\varphi_i(Z)$ a.s. for every i. Formally, this is nothing but the optimality condition for the risk-sharing or infimal convolution problem

$$\inf_{\mathbf{X}\in\mathcal{A}(X)} \sum_{i=1}^{p} \mathbb{E}(\psi_i(X_i)), \tag{4.1}$$

where $\psi_i = \varphi_i^*$ (the Legendre Transform of φ_i). This suggests a definition of comonotone allocations as the allocations that solve a risk-sharing problem of the type above. This has a natural interpretation in terms of risk-sharing, but one has to be cautious about such a definition whenever the ψ_i functions are degenerate[2]. Indeed, if all the ψ_i functions are constant, then any allocation is comonotone in that sense! This means that one has to impose strict convexity in the definition. We shall actually go one step further in *quantifying* strict convexity as follows. Given an arbitrary collection $w = (w_1, ..., w_p)$ of strictly convex functions, we will say that an allocation is w-strictly comonotone whenever it solves a risk-sharing problem of the form (4.1) for some ψ_i functions which are *more convex* than the w_i (i.e. $\psi_i - w_i$ is convex for every i). Allocations which can be approached (in law) by strictly w-comonotone will be called comonotone. Since they solve a strictly convex risk-sharing problem, w-strictly comonotone allocations are efficient and the main goal of this section will be to generalize the univariate comonotone dominance result. We shall indeed prove that for any allocation $\mathbf{X} \in \mathcal{A}(X)$ and any

[2]In the univariate case, the situation is much simpler since one can take $Z = X$, and since the X_i variables sum up to X, each convex function φ_i has to be differentiable i.e. all the ψ_i necessarily are strictly convex. In other words, degeneracies can be ruled out easily in the univariate case.

choice of w, there is a w-comonotone allocation $\mathbf{Y} \in \mathcal{A}(X)$ that dominates $\mathbf{X}$ (strictly whenever $\mathbf{X}$ is not itself w-comonotone). The full proof is detailed in Section 6, but its starting point is quite intuitive and consists of studying the optimization problem:

$$\inf \Big\{ \sum_{i=1}^{p} \mathbb{E}(w_i(Y_i)) \ : \ (Y_1, ..., Y_p) \in \mathcal{A}(X), \ Y_i \succcurlyeq X_i, \ i = 1, ..., p \Big\}. \tag{4.2}$$

Clearly, the solution $\mathbf{Y}$ of (4.2) dominates $\mathbf{X}$. A careful study of the dual of (4.2) will enable us to prove that $\mathbf{Y}$ is necessarily w-comonotone, thus giving the desired multivariate extension of Landsberger and Meilijson's comonotone dominance principle. Note also, that our proof is constructive since it relies on an explicit (although difficult to solve in practice) convex minimization problem.

This section is organized as follows. In paragraph 4.1, we shall reformulate the problem in terms of joint laws rather than random allocations. This is purely technical but will enable us to gain some linearity and some compactness in (4.2). We then define precisely our concepts of multivariate comonotonicity in paragraph 4.2. Paragraph 4.3 states the multivariate comonotone dominance result, i.e. the multivariate generalization of Landsberger and Meilijson's results. Finally, in paragraph 4.4, we gather several remarks on multivariate comonotonicity and emphasize some important qualitative differences between the univariate and multivariate cases.

4.1 From random vectors to joint laws

From now on, it is assumed that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic, that there are p agents and that risk is d-dimensional. X is a given $\mathbb{R}^d$-valued L^∞ random vector modeling an aggregate random multivariate risk, while $\mathbf{X} = (X_1, ..., X_p)$ is a given L^∞ sharing of X among the p agents, that is

$$X = \sum_{i=1}^{p} X_i.$$

Let $\gamma_0 := \mathcal{L}(\mathbf{X})$ be the joint law of $\mathbf{X}$ and $m_0 := \mathcal{L}(X)$. Let γ be a probability measure on $(\mathbb{R}^d)^p$ and γ^i denote its i-th marginal. Note that, $\mathcal{L}(Y_i)$ is the i-th marginal of $\mathcal{L}(\mathbf{Y})$. Let $\Pi_\Sigma \gamma$ be the probability measure on $\mathbb{R}^d$ defined by

$$\int_{\mathbb{R}^d} \varphi(z) d\Pi_\Sigma \gamma(z) = \int_{\mathbb{R}^{d\times p}} \varphi(\sum_{i=1}^{p} x_i) d\gamma(x_1, ..., x_p), \ \forall \varphi \in C_0(\mathbb{R}^d, \mathbb{R}), \tag{4.3}$$

(where C_0 denotes the space of continuous functions that tend to 0 at ∞). It follows from this definition that if $\gamma = \mathcal{L}(\mathbf{Y})$, then $\Pi_\Sigma \gamma = \mathcal{L}(\sum Y_i)$. Hence, if $\mathbf{Y} \in \mathcal{A}(X)$ and $\gamma = \mathcal{L}(\mathbf{Y})$, then $\Pi_\Sigma \gamma = m_0 = \mathcal{L}(X)$. In other words, if $\gamma = \mathcal{L}(\mathbf{Y})$ with $\mathbf{Y} \in \mathcal{A}(X)$, then

$$\int \varphi(x_1 + ... + x_d) d\gamma(x_1, ..., x_d) = \int \varphi(z) dm_0(z), \ \forall \varphi \in C_0(\mathbb{R}^d, \mathbb{R}). \quad (4.4)$$

Since $\mathbf{Y}$ is bounded, γ is compactly supported. It follows from the next lemma that $\{\mathcal{L}(\mathbf{Y}), \ \mathbf{Y} \in \mathcal{A}(X)\}$ coincides with the set of compactly supported probability measures γ on $(\mathbb{R}^d)^p$ that satisfy (4.4):

Lemma 4.1. *Assume $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic. If γ is a compactly supported probability measure on $(\mathbb{R}^d)^p$ and satisfies (4.4), then there exists a random vector $\mathbf{Y} = (Y_1, ..., Y_p) \in \mathcal{A}(X)$ such that $\mathcal{L}(\mathbf{Y}) = \gamma$. Hence $\{\mathcal{L}(\mathbf{Y}), \ \mathbf{Y} \in \mathcal{A}(X)\} = \mathcal{M}(m_0)$, where $\mathcal{M}(m_0)$ is the set of compactly supported probability measures on $(\mathbb{R}^d)^p$ such that $\Pi_\Sigma \gamma = m_0 = \Pi_\Sigma \gamma_0$.*

In the sequel, joint laws $\mathcal{M}(m_0)$ will be used instead of admissible allocations $\mathcal{A}(X)$. For compactness issues, a closed ball $B \in \mathbb{R}^d$ centered at 0 such that m_0 is supported by B^p is chosen, and attention is restricted to the set of elements of $\mathcal{M}(m_0)$ supported by pB (meaning that only risk-sharings of X whose components take value in B will be considered). We thus define

$$\mathcal{M}_B(m_0) := \{\gamma \in \mathcal{M}(m_0) \ : \ \gamma(B^p) = 1\}.$$

4.2 Efficiency and comonotonicity in the multivariate case

Let $\mathcal{C}$ be the cone of convex and continuous functions on B, dominance and efficiency in terms of joint laws are defined as follows:

Definition 4.2. *Let γ and π be in $\mathcal{M}_B(m_0)$, then γ dominates π whenever*

$$\int_{B^p} \sum_{i=1}^{p} \varphi_i(x_i) d\gamma(x_1, ..., x_p) \leq \int_{B^p} \sum_{i=1}^{p} \varphi_i(x_i) d\pi(x_1, ..., x_p) \quad (4.5)$$

for all functions $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$. If, in addition, inequality (4.5) is strict whenever the φ_i functions are further assumed to be strictly convex, then γ is said to dominate strictly π. The allocation $\gamma \in \mathcal{M}_B(m_0)$ is efficient if there is no other allocation in $\mathcal{M}_B(m_0)$ that strictly dominates it.

Given $\gamma_0 \in \mathcal{M}_B(m_0)$, it is easy to check (taking functions $\varphi_i(x) = |x_i|^n$ in (4.5) and letting $n \to \infty$) that any $\gamma \in \mathcal{M}(m_0)$ dominating γ_0 (without the restriction that it is supported on B^p) actually belongs to $\mathcal{M}_B(m_0)$. Hence the choice to only consider allocations supported by B^p is in fact not restrictive. Indeed, if γ is supported by B^p, then efficiency of γ in the usual sense, i.e. without restricting to competitors supported by B^p, is *equivalent* to efficiency among competitors supported by B^p.

To define comonotonicity, let $\psi := (\psi_1, ..., \psi_p)$ be a family of strictly convex continuous functions (defined on B). For any $x \in pB$, let us consider the risk sharing (or infimal convolution) problem:

$$\Box_i \psi_i(x) := \inf \left\{ \sum_{i=1}^p \psi_i(y_i) \ : \ y_i \in B, \ \sum_{i=1}^p y_i = x \right\}.$$

This problem admits a unique solution which will be denoted

$$T_\psi(x) := (T_\psi^1(x), ..., T_\psi^p(x)).$$

Note that, by definition

$$\sum_{i=1}^p T_\psi^i(x) = x, \ \forall x \in pB. \tag{4.6}$$

The map $x \mapsto T_\psi(x)$ gives the optimal way to share x so as to minimize the total cost when each individual cost is ψ_i. It defines the efficient allocation $T_\psi(X) := (T_\psi^1(X), ..., T_\psi^p(X))$ with joint law γ_ψ defined by:

$$\int_{B^p} f(y_1, ..., y_p) d\gamma_\psi(y) := \int_{pB} f(T_\psi(x)) dm_0(x)$$

for any $f \in C(B^p)$. One then defines comonotonicity as follows:

Definition 4.3. *An allocation $\gamma \in \mathcal{M}_B(m_0)$ is strictly comonotone if there exists a family $\psi := (\psi_1, ..., \psi_p)$ of strictly convex continuous functions such that $\gamma = \gamma_\psi$. Given a family $w := (w_1, ..., w_p)$ of strictly convex functions in $C^1(B)$, an allocation $\gamma \in \mathcal{M}_B(m_0)$ is w-strictly comonotone if there exists a family $\psi := (\psi_1, ..., \psi_p)$ of convex continuous functions such that $\psi_i - w_i \in \mathcal{C}$ for every i and $\gamma = \gamma_\psi$.*

We shall soon show that strictly comonotone random vectors are in the image of monotone operators (subgradients of convex functions), evaluated at the same random vector, $p(X)$, which justifies the terminology "comonotonicity" in the multivariate setting. By definition, any strictly comonotone allocation is efficient. As the set of strictly comonotone allocations is not closed, we are led to introduce another definition.

Definition 4.4. *An allocation $\gamma \in \mathcal{M}_B(m_0)$ is comonotone if there exists a sequence of strictly comonotone allocations that weakly star converges to γ. Given a family $w := (w_1, ..., w_p)$ of strictly convex functions in $C^1(B)$, an allocation $\gamma \in \mathcal{M}_B(m_0)$ is w-comonotone, if there exists a sequence of w-strictly comonotone allocations that weakly star converges to γ.*

Definitions 4.3 and 4.4 will be discussed in more detail in paragraph 4.4. To understand the previous notions of comonotonicity and in particular why these allocations are called comonotone, it is important to understand the structure of the T_ψ maps.

Let us first ignore regularity issues and further assume that the ψ_i functions are smooth as well as ψ_i^* their Legendre transforms. Without the constraints $x_i \in B$, then the optimality conditions imply that there is some multiplier $p = p(x)$ such that

$$\nabla \psi_i(T_\psi^i(x)) = p, \text{ hence, } T_\psi^i(x) = \nabla \psi_i^*(p).$$

Using (4.6), one gets

$$x = \sum_{j=1}^{p} \nabla \psi_j^*(p), \text{ hence, } p = \nabla (\sum_{j=1}^{p} \psi_j^*)^*(x),$$

thus,

$$T_\psi^i(x) = \nabla \psi_i^* \left(\nabla (\sum_{j=1}^{p} \psi_j^*)^*(x) \right).$$

The maps T_ψ^i are therefore composed of gradients of convex functions that sum up to the identity. In dimension 1, gradients of convex functions are simply monotone maps (and so are composed of such maps). In higher dimensions, a richer and more complicated structure emerges that will be discussed later. Let us now consider the full problem with the constraints that $x_i \in B$ and still assume that the ψ_i functions are smooth, then the optimality conditions read as the existence of a p and a $\lambda_i \geq 0$ such that $\nabla \psi_i(T_\psi^i(x)) = p - \lambda_i T_\psi^i(x)$ holds together with the complementary slackness conditions: $\lambda_i = 0$ whenever $T_\psi^i(x)$ lies in the interior of B.

4.3 A multivariate dominance result and equivalence between efficiency and comonotonicity

Let us fix an allocation $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ such that $\mathbf{X} \in B^p$ a.s., and set $\gamma_0 = \mathcal{L}(\mathbf{X})$ so that $\gamma_0 \in \mathcal{M}_B(m_0)$. A family $w := (w_1, ..., w_p)$ of C^1

functions is also given, each of them being strictly convex on B as in section 4.2. The first main result in the multivariate case is a dominance result, it states that every allocation is dominated by a w-comonotone one and that the dominance is strict if the initial allocation is not itself w-comonotone.

Theorem 4.5. *Let $\gamma_0 = \mathcal{L}(\mathbf{X})$ and w be as above. Then there exists some $\gamma \in \mathcal{M}(m_0)$ that is w-comonotone and dominates γ_0. Moreover if γ_0 is not itself w-comonotone, then γ strictly dominates γ_0.*

The proof of this result will be given in section 6. Without giving details at this point, let us explain the main arguments of the proof:

- The optimization problem (4.2) admits a unique solution $\mathbf{Y}$ with law $\gamma = \mathcal{L}(\mathbf{Y})$, which is efficient and dominates $\gamma_0 = \mathcal{L}(\mathbf{X})$.
- One then proves that γ is necessarily w-comonotone, by showing that that w-comonotonicity is an optimality condition for (4.2). As usual in convex programming, optimality conditions can be obtained by duality. This leads to consider the problem

$$\inf \Big\{ \mathbb{E}\Big(\sum_{i=1}^{p} \psi_i(X_i) - \Box_i \psi_i (\sum_{i=1}^{p} X_i) \Big) \ : \ \psi_i - w_i \text{ convex, } \forall i \Big\}. \quad (4.7)$$

  By a careful study of (4.7), one can prove (but this is rather technical) that γ is w-comonotone.
- It remains to show that γ strictly dominates γ_0 unless γ_0 is itself w-comonotone. From lemma 2.2, it suffices to show that $\mathbf{Y} \neq \mathbf{X}$. But if γ_0 is not w-comonotone, then $\mathbf{X}$ cannot be optimal for (4.2) and thus $\mathbf{Y} \neq \mathbf{X}$.

In terms of efficiency, the following thus holds:

Theorem 4.6. *Let $\gamma \in \mathcal{M}_B(m_0)$ and w be as before. Then*

1. *if γ is strictly w-comonotone, then it is efficient,*
2. *if γ is efficient, then it is w-comonotone for any w,*
3. *the closure for the weak-star topology of efficient allocations coincides with the set of w-comonotone allocations (which is therefore independent of w).*

Proof. Point 1 is a property already mentioned several times. Point 2 follows from Theorem 4.5 and point 3 follows from points 1 and 2.

□

Note that by definition, if γ_0 is strictly w-comonotone then the value of problem (4.7) is zero. We shall also prove (see section 6) weak form of the converse, namely that if the value of problem (4.7) is zero then γ_0 is w-comonotone. Therefore, the value of (4.7) as a function of the joint law γ_0 can be viewed as a numerical criterion for comonotonicity and thus for efficiency. One can therefore, in principle, use on data this value as a test statistic for efficiency.

4.4 Remarks on multivariate comonotonicity

Comparison with the notion of μ-comonotonicity of [20]. The notion of multivariate comonotonicity considered in this paper is to be related to the notion of μ-comonotonicity proposed by Ekeland, Galichon and Henry in [20]. Recall the alternative characterization of comonotonicity given in the univariate case in Proposition 2.5: X_1 and X_2 are comonotone if and only if $\varrho_\mu(X_1+X_2) = \varrho_\mu(X_1) + \varrho_\mu(X_2)$ for a measure μ that is sufficiently regular. In dimension d, [20] have introduced the concept of μ-comonotonicity, based on this idea: if μ is a probability measure on $\mathbb{R}^d$ which does not give positive mass to small sets, two random vectors X_1 and X_2 on $\mathbb{R}^d$ are called *μ-comonotone* if and only if

$$\varrho_\mu(X_1+X_2) = \varrho_\mu(X_1) + \varrho_\mu(X_2),$$

where the (multivariate) *maximum correlation functional* (see e.g. [35] or [20]) is defined by

$$\varrho_\mu(X) = \sup_{\tilde{Y}\sim\mu} \mathbb{E}\left(X\cdot\tilde{Y}\right).$$

The authors of [20] show that X_1 and X_2 are μ-comonotone if and only if there are two convex functions ψ_1 and ψ_2, and a random vector $U \sim \mu$ such that

$$X_1 = \nabla\psi_1(U) \text{ and } X_2 = \nabla\psi_2(U)$$

holds almost surely. Therefore, the present notion of multivariate comonotonicity approximately consists of calling X_1 and X_2 comonotone if and only if there is some measure μ such that X_1 and X_2 are μ-comonotone. There are, however, qualifications to be added. Indeed, [20] require some regularity on the measure μ. In the current setting, no regularity restrictions are imposed on μ; but instead restrictions on the convexity of ψ_1 and ψ_2 have to be imposed to define the notion of w-comonotonicity before passing to the limit. Although not equivalent, these two sets of restrictions originate from the same concern: two random vectors are always optimally coupled with very degenerate distributions, such as the distribution of constant vectors. Therefore one needs to exclude these degenerate cases in order to avoid a definition which would be void of substance. This is the very reason why the strictly convex w_i functions had to be introduced.

Comonotone allocations do not form a bounded set. In the scalar case, comonotone allocations are parameterized by the set of nondecreasing functions summing to the identity map. This set of functions is convex and equilipschitz hence compact (up to adding constants summing up to zero). This compactness is no longer true in higher dimensions (at least when $w = 0$ and we work on the whole space instead of B), and we believe that this is a major structural difference with respect to the univariate case. For simplicity assume that $p = 2$. As outlined in paragraph 4.2, a comonotone allocation (X_1, X_2) of X is given by a pair of functions that are composed of gradients of convex functions and sum up to the identity map. It is no longer true, in dimension 2 that this set of maps is compact (up to constants). Indeed, let us take $n \in \mathbb{N}^*$, and quadratic ψ_1 and ψ_2 of the form

$$\psi_i(x) = \frac{1}{2}\left\langle S_i^{-1}x, x\right\rangle,\ i = 1, 2,\ x \in \mathbb{R}^2$$

with

$$S_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{8\sqrt{n}} \\ \frac{1}{8\sqrt{n}} & \frac{1}{2n} \end{pmatrix}, \qquad S_2 = \begin{pmatrix} \frac{1}{2} & \frac{-1}{8\sqrt{n}} \\ \frac{-1}{8\sqrt{n}} & \frac{1}{2n} \end{pmatrix}.$$

Then the corresponding map T_ψ is linear, and T_ψ^1 is given by the matrix

$$S_1(S_1 + S_2)^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{n}}{8} \\ \frac{1}{8\sqrt{n}} & \frac{1}{2} \end{pmatrix}$$

which is unbounded.

Comonotone allocations do not form a convex set. Another difference with the univariate case is that the set of maps of the form T_ψ used to define comonotonicity is not convex. To see this (again in the case $p = d = 2$), it is enough to show that the set of pairs of 2×2 matrices

$$K := (S_1(S_1{+}S_2)^{-1}, S_2(S_1{+}S_2)^{-1}),\ S_i \text{ symmetric, positive definite},\ i = 1, 2\}$$

is not convex. First let us remark that if $(M_1, M_2) \in K$ then M_1 and M_2 have a positive determinant. Now for $n \in \mathbb{N}^*$, and $\varepsilon \in (0, 1)$ consider

$$S_1 = \begin{pmatrix} 1 & \sqrt{1-\varepsilon} \\ \sqrt{1-\varepsilon} & 1 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} 1 & -\sqrt{1-\varepsilon} \\ -\sqrt{1-\varepsilon} & 1 \end{pmatrix},$$

$$S_1' = \begin{pmatrix} 1 & \sqrt{n-\varepsilon} \\ \sqrt{n-\varepsilon} & n \end{pmatrix}, \qquad S_2' = \begin{pmatrix} 1 & -\sqrt{n-\varepsilon} \\ -\sqrt{n-\varepsilon} & n \end{pmatrix},$$

and define:

$$M_i = S_i(S_1 + S_2)^{-1},\ M_i' = S_i'(S_1' + S_2')^{-1},\ i = 1, 2.$$

If K was convex then the following matrix would have a positive determinant

$$M_1 + M_1' = \begin{pmatrix} 1 & \frac{\sqrt{1-\varepsilon}}{2} + \frac{\sqrt{n-\varepsilon}}{2n} \\ \frac{\sqrt{1-\varepsilon}}{2} + \frac{\sqrt{n-\varepsilon}}{2} & 1 \end{pmatrix},$$

which is obviously false for n large enough and ε small enough.

5 Concluding remarks

In this paper, we have extended Landsberger and Meilijson's comonotone dominance principle to the multivariate case by introducing the variational problem (4.2). We have then extended the univariate theory of efficient risk-sharing to the case of several goods without perfect substitutability, and we derived tractable implications. Two observations can be made at this point. In the first place, this paper demonstrates *the intrinsic difficulty of the multivariate case,* as many features of the univariate case do not extend to higher dimensions: computational ease, the compactness and convexity of efficient risk-sharing allocations. Second, it illustrates *the need for qualification* inherent to the multivariate case. Contrary to the univariate case, the need to quantify strict convexity as in this paper comes by no coincidence. In fact, just as [20] impose regularity conditions on their "baseline measure" to avoid degeneracy, we work with cones which are strictly included in the cone of convex functions by quantifying the strict convexity of the functions used.

Getting back to our initial motivation, namely, finding testable implications of efficiency for the concave order, we already emphasized in paragraph 4.3 that one obtains as a byproduct of our variational approach a numerical criterion that could in principle be used as a test statistic for comonotonicity and thus for efficiency. We thus believe that the present work paves the way for an interesting research agenda. First of all, an efficient algorithm to decide whether a given allocation in the multivariate case is comonotone or not remains to be discovered – we are currently investigating this point. The convex nature of the underlying optimization problem helps, but the constraints of problem $(\mathcal{P}^*)$ are delicate to handle numerically. Finally, this work opens a research agenda on the empirical relevance of the multivariate theory confronted to the data: do observations of realized allocations of risk satisfy restrictions imposed by multivariate comonotonicity? As mentioned above, tests in the univariate case have been performed by [4] and [37] and suggest rejection. But there is hope that in the more flexible setting of multivariate risks, efficiency would be less strongly rejected.

6 Proofs

6.1 Proof of Lemma 2.2

Clearly $1 \Rightarrow 2$ and $3 \Rightarrow 1$ are obvious. To prove that $2 \Rightarrow 3$, assume that 2 holds true. Let $\mu := \mathcal{L}(X)$ and $\nu := \mathcal{L}(Y)$. These probability measures are supported by some closed ball B, and the Cartier-Fell-Meyer theorem states that there is a measurable family of conditional probability measures $(T_x)_{x\in B}$ such that T_x has mean x and for every f continuous function, one has

$$\mathbb{E}(f(Y)) = \int_B f(y)d\nu(y) = \int_B \int_B f(y)dT_x(y)d\mu(x)$$

Since $\mu \neq \nu$, $\mu(\{x \in B \ : \ T_x \neq \delta_x\}) > 0$, one deduces from Jensen's inequality that for every strictly convex function φ, $\mathbb{E}(\varphi(Y)) > \mathbb{E}(\varphi(X))$.

6.2 Proof of Lemma 4.1

For notational simplicity, assume that $d = 1$, $p = 2$, X takes values in $[0, 2]$ a.s. (so that m_0 has support in $[0, 2]$) and γ is supported by $[0, 1]^2$. For every $n \in \mathbb{N}^*$ and $k \in \{0, ..., 2^{n+1}\}$, set

$$X^n := \sum_{k=0}^{2^{n+1}} \frac{k}{2^n}\mathbf{1}_{A_{k,n}}, \quad \text{where } A_{k,n} := \left\{\omega \in \Omega \ : \ X(\omega) \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[\right\},$$

and

$$C_{k,n} := \left\{(y_1, y_2) \in [0, 1]^2 \ : \ y_1 + y_2 \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[\right\}.$$

Decompose the strip $C_{k,n}$ into a partition by triangles

$$C_{k,n} = \bigcup_{k\leq i+j\leq k+1} T^{i,j}_{k,n},\ T^{i,j}_{k,n} := C_{k,n} \cap \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right[\times \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right[.$$

Since $\Pi_\Sigma(\gamma) = m_0$ one has:

$$\mathbb{P}(A_{k,n}) = \gamma(C_{k,n}) = \sum_{k\leq i+j\leq k+1} \gamma(T^{i,j}_{k,n}),$$

and since $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic, it follows from Lyapunov's convexity theorem (see [29]) that there exists a partition of $A_{k,n}$ into measurable subsets $A_{k,n}^{i,j}$ such that

$$\gamma(T_{k,n}^{i,j}) = \mathbb{P}(A_{k,n}^{i,j}), \forall (i,j) \in \{0, ..., 2^n\} \ : \ k \leq i+j \leq k+1. \tag{6.1}$$

Choose $(y_1, y_2)_{k,n}^{i,j} \in T_{k,n}^{i,j}$ and define

$$\mathbf{Y}^n = (Y_1^n, Y_2^n) := \sum_{k=0}^{2^{n+1}} \sum_{k \leq i+j \leq k+1} (y_1, y_2)_{k,n}^{i,j} \mathbf{1}_{A_{k,n}^{i,j}}.$$

We may also choose inductively the partition of $A_{k,n}$ by the $A_{k,n}^{i,j}$ to be finer and finer with respect to n. By construction, one obtains

$$\max\Big(\|X^n - X\|_{L^\infty}, \|X^n - Y_1^n - Y_2^n\|_{L^\infty}, \|\mathbf{Y}^{n+1} - \mathbf{Y}^n\|_{L^\infty}\Big) \leq \frac{1}{2^n},$$

so that $\mathbf{Y}^n$ is a Cauchy sequence in L^∞, and thus converges to some $\mathbf{Y} = (Y_1, Y_2)$. One then sees that $Y_1 + Y_2 = X$, and passing to the limit in (6.1), it follows that $\mathcal{L}(\mathbf{Y}) = \gamma$.

6.3 Proofs and variational characterization for the multivariate dominance result

The proofs will very much rely on the linear programming problem:

$$(\mathcal{P}^*) \ \sup_{\gamma \in K(\gamma_0)} - \int_{B^p} \sum_{i=1}^{p} w_i(x_i) d\gamma(x)$$

where $K(\gamma_0)$ consists of all $\gamma \in \mathcal{M}_B(m_0)$ such that for each i the marginal γ^i dominates the corresponding marginal of γ_0 i.e.:

$$\int_{B^p} \varphi(x_i) d\gamma(x) \leq \int_{B^p} \varphi(x_i) d\gamma_0(x), \forall \varphi \text{ convex on } B.$$

Problem $(\mathcal{P}^*)$ presents similarities with the problem solved in [22]. In the optimal transport problem considered in [22], one minimizes the average of some quadratic function over joint measures having prescribed marginals whereas $(\mathcal{P}^*)$ includes dominance constraints on the marginals. To shorten notations, define

$$\eta(x) := - \sum_{i=1}^{p} w_i(x_i), \forall x = (x_1, ..., x_p) \in B^p$$

$(\mathcal{P}^*)$ is the dual problem (see the next lemma for details) of

$$(\mathcal{P})\inf\Big\{\int_{B^p}\Big(\sum_{i=1}^p\varphi_i(x_i)-\varphi_0\Big(\sum_{i=1}^p x_i\Big)\Big)d\gamma_0(x),\ (\varphi_0,...,\varphi_p)\in E\Big\},$$

where E consists of all families $\varphi:=(\varphi_1,....,\varphi_p,\varphi_0)\in C(B)^p\times C(pB)$ such that $\varphi_i\in\mathcal{C}$ and

$$\sum_{i=1}^p\varphi_i(x_i)-\varphi_0\Big(\sum_{i=1}^p x_i\Big)\geq-\sum_{i=1}^p w_i(x_i).$$

It will also be convenient to consider

$$(\mathcal{Q})\inf\Big\{J(\psi)\ ,\ \psi=(\psi_1,...,\psi_p):\text{ each }\psi_i\text{ is such that }\psi_i-w_i\text{ is convex}\Big\}$$

with

$$J(\psi):=\int_{B^p}\Big(\sum_{i=1}^p\psi_i(x_i)-\square_i\psi_i\Big(\sum_{i=1}^p x_i\Big)\Big)d\gamma_0(x).$$

Note that by construction $J(\psi)\geq0$ for every admissible ψ and $J(\psi)=0$ if and only if $\gamma_0=\gamma_\psi$.

Lemma 6.1. *The following holds*

$$\max(\mathcal{P}^*)=\inf(\mathcal{P})=\inf(\mathcal{Q})-\int_{B^p}\sum_{i=1}^p w_i(x_i)d\gamma_0(x).$$

Proof. Let us write $(\mathcal{P})$ in the form

$$\inf_{\varphi=(\varphi_1,...,\varphi_p,\varphi_0)\in C(B)^p\times C(pB)}F(\Lambda\varphi)+G(\varphi)$$

where $\Lambda:\,C(B)^p\times C(pB)\to C(B^p)$ is the linear continuous map defined by

$$\Lambda\varphi(x):=\sum_{i=1}^p\varphi_i(x_i)-\varphi_0\Big(\sum_{i=1}^p x_i\Big),\ \forall x=(x_1,...,x_p)\in B^p,$$

and F, G are the convex lower semicontinuous (for the uniform norm) functionals defined by

$$F(\theta)=\begin{cases}\int_{B^p}\theta d\gamma_0\text{ if} & \theta\geq\eta\\ +\infty & \text{otherwise}\end{cases},\ \forall\theta\in C(B^p)$$

$$G(\varphi) = \begin{cases} 0 \text{ if } & (\varphi_1, ..., \varphi_p) \in \mathcal{C}^p \\ +\infty & \text{otherwise} \end{cases}, \forall \varphi = (\varphi_1, ..., \varphi_p, \varphi_0) \in C(B)^p \times C(pB).$$

It is easy to see that the assumptions of Fenchel-Rockafellar's duality theorem (see for instance [18]) are satisfied and thus

$$\inf(\mathcal{P}) = \max_{\gamma \in \mathcal{M}(B^p)} -F^*(\gamma_0 - \gamma) - G^*(\Lambda^*(\gamma - \gamma_0)).$$

The adjoint of Λ, Λ^* is easily computed as : $\mathcal{M}(B^p) \to \mathcal{M}(B)^p \times \mathcal{M}(pB)$ (where $\mathcal{M}$ denotes the space of Radon measures):

$$\Lambda^* \gamma = (\gamma^1, ..., \gamma^p, -\Pi_\Sigma \gamma), \ \forall \gamma \in \mathcal{M}(B^p).$$

Direct computations give

$$F^*(\gamma - \gamma_0) = \begin{cases} -\int_{B^p} \eta d\gamma \text{ if } & \gamma \geq 0 \\ +\infty & \text{otherwise} \end{cases}$$

and

$$G^*(\Lambda^*(\gamma - \gamma_0)) = \begin{cases} 0 \text{ if } & \gamma \in K(\gamma_0) \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore $(\mathcal{P}^*)$ is the dual of $(\mathcal{P})$ in the usual sense of convex programming and $\max(\mathcal{P}^*) = \inf(\mathcal{P})$. To prove that

$$\inf(\mathcal{P}) = \inf(\mathcal{Q}) - \int_{B^p} \sum_{i=1}^p w_i(x_i) d\gamma_0(x),$$

take $\varphi \in E$ and $\psi_i := w_i + \varphi_i$ for $i = 1, .., p$, the constraint then reads as

$$\sum_{i=1}^p \psi_i(x_i) \geq \varphi_0 \Big(\sum_{i=1}^p x_i \Big), \ \forall x \in B^p.$$

Now in $(\mathcal{P})$, one needs to choose φ_0 as large as possible without violating this constraint. Thus the best φ_0 given $(\varphi_1, ..., \varphi_p)$ is

$$\varphi_0 = \square_i \psi_i,$$

which proves the desired identity.

$\square$

Lemma 6.2. *Let ψ_i be such that $\psi_i - w_i \in \mathcal{C}$ for every i and $g = (g_1, ..., g_p) \in \mathcal{C}^p$. Then*

$$\lim_{\delta \to 0^+} \frac{1}{\delta}[J(\psi + \delta g) - J(\psi)] = \sum_{i=1}^{p} \int_B g_i(x_i) d(\gamma_0^i - \gamma_\psi^i)$$
$$= \int_{B^p} \sum_{i=1}^{p} g_i(x_i) d(\gamma_0 - \gamma_\psi)(x).$$

Proof. For $\delta > 0$, one first gets that

$$\frac{1}{\delta}[J(\psi + \delta g) - J(\psi)] = \sum_{i=1}^{p} \int_B g_i(x_i) d(\gamma_0^i) -$$
$$\int_{pB} \frac{1}{\delta}\Big(\Box_i(\psi_i + \delta g_i)(x) - \Box_i \psi_i(x)\Big) dm_0(x).$$

Note that the integrand in the second term is bounded since g is. Now fix some $(x_1, ..., x_p) \in B^p$, and set $x = \sum_{i=1}^p x_i$, $y_i := T_\psi^i(x)$ and $y_i^\delta := T_{\psi+\delta g}(x)$. Since $\sum_{i=1}^p y_i = \sum_{i=1}^p y_i^\delta = x$, it comes as a direct consequence of the definition of infimal convolutions that:

$$\frac{1}{\delta}\Big(\Box_i(\psi_i + \delta g_i)(x) - \Box_i \psi_i(x)\Big) \le \sum_{i=1}^{p} g_i(y_i) \tag{6.2}$$

and

$$\frac{1}{\delta}\Big(\Box_i(\psi_i + \delta g_i)(x) - \Box_i \psi_i(x)\Big) \ge \sum_{i=1}^{p} g_i(y_i^\delta). \tag{6.3}$$

Using the compactness of B and the strict convexity of ψ_i, it is easy to check that $y_i^\delta \to y_i$ as $\delta \to 0^+$. Therefore, from (6.2) and (6.3) one has

$$\lim_{\delta \to 0^+} \frac{1}{\delta}\Big(\Box_i(\psi_i + \delta g_i)(x) - \Box_i \psi_i(x)\Big) = \sum_{i=1}^{p} g_i(T_\psi^i(x))$$

and this holds for every $x \in pB$. It then follows from Lebesgue's dominated convergence theorem that

$$\lim_{\delta \to 0^+} \frac{1}{\delta}[J(\psi + \delta g) - J(\psi)] = \sum_{i=1}^{p} \int_B g_i(x_i) d(\gamma_0^i) - \sum_{i=1}^{p} \int_{pB} g_i(T_\psi^i(x)) dm_0(x)$$
$$= \sum_{i=1}^{p} \int_B g_i(x_i) d(\gamma_0^i - \gamma_\psi^i) = \int_{B^p} \sum_{i=1}^{p} g_i(x_i) d(\gamma_0 - \gamma_\psi)(x).$$

□

It follows from Lemma 6.2 that, if ψ solves $(\mathcal{Q})$, then γ_ψ dominates γ_0. Hence, if one knew that $(\mathcal{Q})$ possesses solutions, the existence of an ω-strictly comonotone allocation dominating γ_0 would directly follow. Unfortunately, it is not necessarily the case that the infimum in $(\mathcal{Q})$ is attained – or at least we haven't been able to prove without additional conditions. The difficulty here comes from the fact that minimizing sequences need not be bounded (see paragraph 4.4). It may be the case that additional regularity assumptions on γ_0 would guarantee existence. In the present paper no such assumption is made, and a different path is chosen to overcome the difficulty by an appeal to Ekeland's variational principle.

Lemma 6.3. *Letting $\varepsilon > 0$, there exists ψ_ε admissible for $(\mathcal{Q})$ such that*

1. $J(\psi_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon$
2. $$\limsup_{\varepsilon \to 0^+} \int_{B^p} \sum_{i=1}^p \varphi_i(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \leq 0$$
   for every $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$
3. $$\liminf_{\varepsilon \to 0^+} \int_{B^p} \sum_{i=1}^p \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq 0$$
   for $\varphi_i^\varepsilon = \psi_{i,\varepsilon} - w_i$ (these are convex functions by definition).

Proof. For $\varepsilon > 0$, let f_ε be admissible for $(\mathcal{Q})$ and such that

$$J(f_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon.$$

Let then $k_\varepsilon > 0$ be such that

$$\lim_{\varepsilon \to 0^+} \varepsilon k_\varepsilon [1 + \|f_\varepsilon\|] = 0 \text{ (for instance } k_\varepsilon = \frac{1}{\varepsilon^{1/2}(1 + \|f_\varepsilon\|)}). \tag{6.4}$$

It follows from Ekeland's variational principle (see [17] and [5]) that for every $\varepsilon > 0$, there is some ψ_ε admissible for $(\mathcal{Q})$ such that

$$\|\psi_\varepsilon - f_\varepsilon\| \leq \frac{1}{k_\varepsilon}, \; J(\psi_\varepsilon) \leq J(f_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon, \tag{6.5}$$

where $\|h\|$ stands for the sum of the uniform norms of the h_i functions, and

$$J(\psi) \geq J(\psi_\varepsilon) - k_\varepsilon \varepsilon \|\psi - \psi_\varepsilon\|, \; \forall \psi = (\psi_1, ..., \psi_p) \; : \; \psi_i - w_i \in \mathcal{C}, \; \forall i. \tag{6.6}$$

Taking $\psi = \psi_\varepsilon + \delta\varphi$ with $\delta > 0$ and $\varphi \in \mathcal{C}^p$ in (6.6), dividing by δ and letting $\delta \to 0^+$, one thus gets by the virtues of Lemma 6.2

$$\int_{B^p} \sum_{i=1}^{p} \varphi_i(x_i) d(\gamma_0 - \gamma_{\psi_\varepsilon}) \geq -k_\varepsilon \varepsilon \|\varphi\|. \tag{6.7}$$

Using (6.4) and letting $\varepsilon \to 0^+$ yields

$$\limsup_{\varepsilon \to 0^+} \int_{B^p} \sum_{i=1}^{p} \varphi_i(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \leq 0 \tag{6.8}$$

for every $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$. To prove the last assertion of the lemma, write $\psi_\varepsilon = \varphi^\varepsilon + w$ with $\varphi^\varepsilon \in \mathcal{C}^p$. Then for $\delta \in (0,1)$ one has $\psi_\varepsilon - \delta\varphi^\varepsilon = (1-\delta)\varphi^\varepsilon + w$, and then (6.6) can be applied to $\psi_\varepsilon - \delta\varphi^\varepsilon$, leading to

$$\frac{1}{\delta}[J(\psi_\varepsilon - \delta\varphi^\varepsilon) - J(\psi_\varepsilon)] \geq -k_\varepsilon \varepsilon \|\varphi^\varepsilon\|,$$

which, letting $\delta \to 0^+$ and using the same argument as in lemma 6.2, leads in turn to

$$\int_{B^p} \sum_{i=1}^{p} \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq -k_\varepsilon \varepsilon \|\varphi^\varepsilon\|.$$

By (6.4) and (6.5), it follows that

$$k_\varepsilon \varepsilon \|\varphi^\varepsilon\| \leq k_\varepsilon \varepsilon (\|w\| + \|\psi_\varepsilon - f_\varepsilon\| + \|f_\varepsilon\|) \leq k_\varepsilon \varepsilon \|w\| + \varepsilon + k_\varepsilon \varepsilon \|f_\varepsilon\| \to 0 \text{ as } \varepsilon \to 0^+,$$

which enables us to conclude that

$$\liminf_{\varepsilon \to 0^+} \int_{B^p} \sum_{i=1}^{p} \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq 0. \tag{6.9}$$

□

Lemma 6.4. *Let ψ_ε be as in lemma 6.3 and set $\gamma_\varepsilon := \gamma_{\psi_\varepsilon}$ then up to some subsequence, γ_ε weakly star converges to some γ (w-comonotone by construction) such that $\gamma \in \mathcal{M}_B(m_0)$ and γ dominates γ_0. Moreover γ solves $(\mathcal{P}^*)$.*

Proof. By the Banach-Alaoglu-Bourbaki theorem, one may indeed assume that γ_ε weakly star converges to some γ. Obviously, γ is w-comonotone and $\Pi_\Sigma \gamma = \Pi_\Sigma \gamma_0 = m_0$, hence $\gamma \in \mathcal{M}_B(m_0)$. The fact that γ dominates γ_0 is obtained by letting $\varepsilon \to 0^+$ in (6.8). It remains to prove that γ solves $(\mathcal{P}^*)$. Defining $\varphi^\varepsilon := \psi_\varepsilon - w$ as in Lemma 6.3, one has

$$J(\psi_\varepsilon) = \int_{B^p} \sum_{i=1}^p \varphi_i^\varepsilon(x_i) d(\gamma_0 - \gamma_\varepsilon) + \int_{B^p} \eta d(\gamma_\varepsilon - \gamma_0) \to \inf(\mathcal{Q}) \text{ as } \varepsilon \to 0^+.$$

By (6.9), passing to the limit thus yields

$$\inf(\mathcal{Q}) \leq \int_{B^p} \eta d(\gamma - \gamma_0),$$

which, combined with Lemma 6.1, gives

$$\int_{B^p} \eta d\gamma \geq \inf(\mathcal{Q}) + \int_{B^p} \eta d\gamma_0 = \max(\mathcal{P}^*).$$

□

Lemma 6.5. *Let γ be as in lemma 6.4. Then:*

1. *if γ_0 solves $(\mathcal{P}^*)$ then γ_0 is w-comonotone,*
2. *γ strictly dominates γ_0 unless γ_0 is itself w-comonotone.*

Proof. If γ_0 solves $(\mathcal{P}^*)$, it follows from Lemma 6.1 that $\inf(\mathcal{Q}) = 0$. For any minimizing sequence ψ_ε (not necessarily the one constructed in Lemma 6.3) of $(\mathcal{Q})$, the following holds:

$$\begin{aligned} 0 = \lim_{\varepsilon \to 0^+} J(\psi_\varepsilon) &= \lim_{\varepsilon \to 0^+} \int_{B^p} \Big(\sum_{i=1}^p \psi_{i,\varepsilon}(x_i) - \square_i \psi_{i,\varepsilon}\Big(\sum_{i=1}^p x_i \Big)\Big) d\gamma_0(x) \\ &= \lim_{\varepsilon \to 0^+} \int_{B^p} \Big(\sum_{i=1}^p \psi_{i,\varepsilon}(x_i) - \sum_{i=1}^p \psi_{i,\varepsilon}\Big(T^i_{\psi_\varepsilon}\Big(\sum_{i=1}^p x_i \Big)\Big)\Big) d\gamma_0(x). \end{aligned}$$

By a density argument, one may consider ψ_ε a minimizing sequence such that each ψ_ε belongs to $C^1(B)$. Fix $(x_1, ..., x_p)$ and set $x := \sum_{i=1}^p x_i$. Then $y^\varepsilon := T_{\psi_\varepsilon}(x)$ can be characterized by the fact that there is a vector $p \in \mathbb{R}^d$ and a vector (λ_i) of nonnegative weights such that

$$\nabla \psi_{i,\varepsilon}(y_i^\varepsilon) = p - \lambda_i y_i^\varepsilon, \ \lambda_i = 0 \text{ if } y_i^\varepsilon \notin \partial B, \ \sum_{i=1}^p y_i^\varepsilon = x. \tag{6.10}$$

On the other hand, since w_i is strictly convex and $\psi_{i,\varepsilon} - w_i \in \mathcal{C}$, it follows that for any a and b in B^2,

$$\psi_{i,\varepsilon}(b) - \psi_{i,\varepsilon}(a) \geq \nabla\psi_{i,\varepsilon}(a) \cdot (b-a) + \theta_i(|b-a|) \tag{6.11}$$

where function θ_i is defined by

$$\theta_i(t) := \inf\{w_i(b) - w_i(a) - \nabla w_i(a) \cdot (b-a),\ (a,b) \in B^2,\ |a-b| \geq t\}$$

for any $t \in [0, \mathrm{diam}(B)]$. Function θ_i (the modulus of strict convexity of w_i) is a nondecreasing function such that $\theta_i(0) = 0$ and $\theta_i(t) > 0$ for $t > 0$. Combining (6.10) and (6.11) yields

$$\begin{aligned}\sum_{i=1}^p \psi_{i,\varepsilon}(x_i) - \sum_{i=1}^p \psi_{i,\varepsilon}(y_i^\varepsilon) &\geq \sum_{i=1}^p \nabla\psi_{i,\varepsilon}(y_i^\varepsilon) \cdot (x_i - y_i^\varepsilon) + \sum_{i=1}^p \theta_i(|x_i - y_i^\varepsilon|) \\ &= p \cdot \sum_{i=1}^p (x_i - y_i^\varepsilon) - \sum_{i=1}^p \lambda_i y_i^\varepsilon (x_i - y_i^\varepsilon) + \sum_{i=1}^p \theta_i(|x_i - y_i^\varepsilon|) \\ &\geq \sum_{i=1}^p \theta_i(|x_i - y_i^\varepsilon|).\end{aligned}$$

Hence the fact that $J(\psi_\varepsilon) \to 0$ as $\varepsilon \to 0^+$ implies

$$\lim_{\varepsilon \to 0^+} \int_{B^p} \sum_{i=1}^p \theta_i(|x_i - T^i_{\psi_\varepsilon}(\sum_j x_j)|) d\gamma_0(x) = 0$$

so that

$$T_{\psi_\varepsilon}(\sum_j x_j) - x \to 0 \text{ as } \varepsilon \to 0^+ \text{ for } \gamma_0\text{-a.e. } x.$$

Therefore, by Lebesgue's dominated convergence theorem,

$$\begin{aligned}\int_{B^p} f(x) d\gamma_0(x) &= \lim_{\varepsilon \to 0^+} \int_{B^p} f(T_{\psi_\varepsilon}(\sum_j x_j)) d\gamma_0(x) = \lim_{\varepsilon \to 0^+} \int_{pB} f(T_{\psi_\varepsilon}(x)) dm_0(x) \\ &= \lim_{\varepsilon \to 0^+} \int_{B^p} f d\gamma_{\psi_\varepsilon}\end{aligned}$$

holds for all $f \in C(B^p)$. Hence, $\gamma_{\psi_\varepsilon}$ weakly star converges to γ_0 which proves that γ_0 is w-comonotone and Point 1 is proven. We now prove Point 2. If γ_0 is not w-comonotone then by Point 1, it does not solve $(\mathcal{P}^*)$ and thus $\int \eta d(\gamma - \gamma_0) > 0$ so that

$$\int_{B^p} \sum_{i=1}^p w_i(x_i) d\gamma < \int_{B^p} \sum_{i=1}^p w_i(x_i) d\gamma_0$$

hence γ strictly dominates γ_0. This completes the proof.

$\square$

6.4 Proof of Theorem 3.1

Let $\gamma_0 := \mathcal{L}(\mathbf{X}) = \mathcal{L}(X_1, ..., X_p)$ and consider again the minimization problem $(\mathcal{P}^*)$. It follows from Lemma 6.4 that there exists a w-comonotone solution γ to $(\mathcal{P}^*)$. By construction, γ dominates γ_0, and by the definition of w-comontonicity, there is a sequence γ_n of w-strictly comonotone allocations that weakly star converges to γ. Lemma 4.1 allows to write $\gamma_n = \mathcal{L}(\mathbf{Y}_n)$ for some $\mathbf{Y}_n \in \mathcal{A}(X)$ which is obviously comonotone in the univariate sense according to Definition 2.4. Since $\mathbf{Y}_n$ is bounded in L^∞ and each $\mathbf{Y}_n$ is a 1-Lipschitz function of X, one may assume that $\mathbf{Y}_n$ converges uniformly (up to a subsequence) to some $\mathbf{Y} \in \mathcal{A}(X)$. It is obvious that $\mathbf{Y}$ is also comonotone and that $\gamma = \mathcal{L}(\mathbf{Y})$, hence $\mathbf{Y}$ dominates $\mathbf{X}$. The strict dominance assertion follows from Lemma 6.5.

References

[1] K.J. Arrow, Uncertainty and the welfare of medical care, Amer. Econom. Rev. 53 (1963), 941–973.

[2] K.J. Arrow, Optimal insurance and generalized deductibles, Scand. Actuar. J. (1974), 1–42.

[3] A.B. Atkinson, On the measurement of Inequality, Journal of Economic Theory 2 (1970), 244–263.

[4] Attanasio, O., Davis, S., Relative Wage Movements and the Distribution of Consumption, Journal of Political Economy, 104(6), 1227-62.

[5] J.-P. Aubin, I. Ekeland, Applied Nonlinear Analysis, John Wiley ans Sons, New-York, 1984.

[6] K. Borch, Equilibrium in a reinsurance market, Econometrica, 30 (1962), 424–444.

[7] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. 44 (1991), 375–417.

[8] Brown, D.J. and R.L. Matzkin, Testable Restrictions on the Equilibrium Manifold, Econometrica, Vol. 64 (1996), No. 6, 1249-1262.

[9] G. Carlier, R.-A. Dana, Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints, Statistics and Decisions, 24 (2006), 127–152.

[10] G. Carlier, R.-A. Dana, Two-Persons Efficient Risk-Sharing and Equilibria for Concave Law-Invariant Utilities, Economic Theory, 36 (2008), 189–223.

[11] P. Cartier, J.M.G. Fell, P.-A. Meyer, Comparaison des mesures portées par un ensemble convexe compact, Bull. Soc. Math. France, 92 (1964), 435–445.

[12] S.H. Chew, I. Zilcha, Invariance of the efficient sets when the expected utility hypothesis is relaxed, Journal of Economic Behaviour and Organisation 13 (1990) 125–131.

[13] R.-A. Dana, Market behavior when preferences are generated by second order stochastic dominance, Journal of Mathematical Economics 40, (2004), 619-639.

[14] R.-A. Dana, I.I Meilijson, Modelling Agents' Preferences in Complete Markets by Second Order Stochastic Dominance. Working Paper 0333, Cahiers du CEREMADE

[15] D. Denneberg, Non-additive Measures and Integral, Kluwer Academic Publishers, Holland, 1994.

[16] P. Dybvig, Distributional Analysis of Portfolio Choice, Journal of Business 61 (1988) 369–393.

[17] I. Ekeland, On the variational principle. J. Math. Anal. Appl. 47 (1974), 324–353.

[18] I. Ekeland, R. Temam, Convex analysis and variational problems, North-Holland, Amsterdam-Oxford, 1976.

[19] A. Galichon, M. Henry, Dual theory of choice with multivariate risks, forthcoming in the Journal of Economic Theory.

[20] I. Ekeland, A. Galichon, M. Henry, Comonotonic measures of multivariate risk, forthcoming in Mathematical Finance.

[21] H. Föllmer, A. Schied, Stochastic finance. An introduction in discrete time , De Gruyter editor, Berlin 2004.

[22] W. Gangbo, A. Święch, Optimal maps for the multidimensional Monge-Kantorovich problem, Comm. Pure Appl. Math., 51 (1998), 23–45.

[23] E. Jouini, C. Napp, Comonotonic processes, Insurance Math. Econom. 32 (2003), 255–265.

[24] E. Jouini, C. Napp, Conditional comonotonicity, Decis. Econ. Finance 27 (2004), 153–166.

[25] E. Jouini , H. Kallal, (2000) Efficient Trading Strategies in the Presence of Market Frictions, Review of Financial Studies 14 (2000), 343–369.

[26] E. Jouini, W. Schachermayer, N. Touzi, Optimal risk sharing for law invariant monetary utility functions, Math. Finance 18 (2008), 269–292.

[27] M. Landsberger, I.I. Meilijson, Comonotone allocations, Bickel Lehmann dispersion and the Arrow-Pratt measure of risk aversion, Annals of Operation Research 52 (1994), 97–106.

[28] S.F. LeRoy, J. Werner, Principles of Financial Economics. Cambridge University Press, Cambridge, 2001.

[29] A. Lyapunov, Sur les fonctions-vecteurs completement additives, Bull. Acad. Sci. URSS (6) (1940), 465-478.

[30] M. Ludkovski, L. Rüschendorf, On comonotonicity of Pareto optimal risk sharing, Statistics and Probability Letters 78 (2008), 1181-1188.

[31] A. Müller, D. Stoyan, Comparison Methods for Stochastic Models and Risks, Wiley, New-York, 2002.

[32] B. Peleg, M.E. Yaari, A Price Characterisation of Efficient Random Variables, Econometrica 43 (1975), 283–292.

[33] G. Puccetti, M. Scarsini, Multivariate comonotonicity, Journal of Multivariate Analysis 101 (2010), 291–304.

[34] M. Rothschild, J. E. Stiglitz,. Increasing Risk, I. A Definition, Journal of Economic Theory 2 (1970), 225–243.

[35] L. Rüschendorf, Law invariant convex risk measures for portfolio vectors, Statist. Decisions 24 (2006), 97–108.

[36] V. Strassen, The existence of probability measures with given marginals, Ann. Math. Statist. 36 (1965), 423–439.

[37] R.M. Townsend, Risk and Insurance in village India, Econometrica 62 (1994), 539-592.

[38] R. Wilson, The theory of syndicates, Econometrica 55 (1968), 95-115.